Learn R: Mean, Median, and Mode

Median of a Dataset

The median of a dataset is the value that, assuming the dataset is ordered from smallest to largest, falls in the middle. If there are an even number of values in a dataset, the middle two values are the median.
Say we have a dataset with the following ten numbers:

`24, 16, 30, 10, 12, 28, 38, 2, 4, 36`

We can order this dataset from smallest to largest:

`2, 4, 10, 12, 16, 24, 28, 30, 36, 38`

The medians of this dataset are 16 and 24, because they are the fifth- and sixth-positioned observations in this dataset. In other words, there are four observations to the left of 16, and four observations to the right of 24.

If we added another value (say, 28) near the middle of this dataset:

`2, 4, 10, 12, 16, 24, 28, 28, 30, 36, 38`

The new median is equal to 24, because there are 5 values smaller than it, and 5 values larger than it.

The median() Function in R

In R, the median of a vector is calculated using the $median()$ function. The function accepts a vector as an input. If there are an odd number of values in the vector, the function returns the middle value. If there are an even number of values in the vector, the function returns the average of the two medians.

Even:

```
b <- c(3,4,5,6,12)
median(b)
```

The code above outputs 5 as the median, because it is the middle number in the array.

Odd:

```
a <- c(3,4,5,12)
median(a)
```

The code above outputs 4.5, because it takes the average of the two medians, 4 and 5.

Mean of a Dataset

The *mean*, or average, of a dataset is calculated by adding all the values in the dataset and then dividing by the number of values in the set.

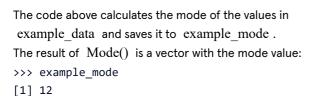For example, for the dataset $[1,2,3]$, the mean is $1+2+3$ $/$ 3 $=$ 2.

The mode() Function in R

In R, the mode of a vector can be calculated using the Mode() function in the DescTools package. The function accepts a vector as an input and returns the most frequently occurring observation in the dataset.

One Mode

```
library(DescTools)

example_data <- c(24, 16, 12, 10,

example_mode <- Mode(example_data
```

◀ ▬▬▬▬▬▬ ▶

The code above calculates the mode of the values in example_data and saves it to example_mode.
The result of Mode() is a vector with the mode value:
```
>>> example_mode
[1] 12
```

Two Modes

If there are multiple modes, the Mode() function will return them as a vector.
Let's look at a vector with two modes, 12 and 24:

```
example_data = c(24, 16, 12, 10,

example_mode = Mode(example_data)
```

◀ ▬▬▬▬ ▶

The result is:

```
>>> example_mode
[1] 12 24
```

The mean() Function

In R, the mean of a vector is calculated using the
mean() function. The function accepts a vector as
input, and returns the average as a numeric.
The code below is used to create a vector and calculate
its mean:

```
a <- c(3,4,5,6)

mean(a)
```

This code outputs the average value of the array
c(3,4,5,6) :
4.5

↓ Print   ⊷ Share ▼